Recall: Homotopy Continuation

- Given F, a polynomial system of equations

$$
\begin{aligned}
f_1(x_1, \ldots, x_n) &= 0 \\
f_2(x_1, \ldots, x_n) &= 0 \\
&\vdots \\
f_n(x_1, \ldots, x_n) &= 0.
\end{aligned}
$$

Recall: Homotopy Continuation

- Given F, a polynomial system of equations

$$f_1(x_1, \dots, x_n) = 0$$

$$f_2(x_1, \dots, x_n) = 0$$

$$\vdots$$

$$f_n(x_1, \dots, x_n) = 0.$$

- Choose and solve instead an (easier) polynomial system G based on characteristics of F.

Recall: Homotopy Continuation

- Given F, a polynomial system of equations

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0 \\ f_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, \ldots, x_n) &= 0. \end{aligned}$$

- Choose and solve instead an (easier) polynomial system G based on characteristics of F.
- Form the homotopy system $H(x, t) = (1 - t) \cdot F(x) + t \cdot G(x)$

Recall: Homotopy Continuation

- Given F, a polynomial system of equations

$$
\begin{aligned}
f_1(x_1, \ldots, x_n) &= 0 \\
f_2(x_1, \ldots, x_n) &= 0 \\
&\vdots \\
f_n(x_1, \ldots, x_n) &= 0.
\end{aligned}
$$

- Choose and solve instead an (easier) polynomial system G based on characteristics of F.
- Form the homotopy system $H(x,t) = (1-t) \cdot F(x) + t \cdot G(x)$
- Use predictor-corrector methods to track the numerical solutions as t moves from $t = 1$ to $t = 0$.

Homotopy Tracking

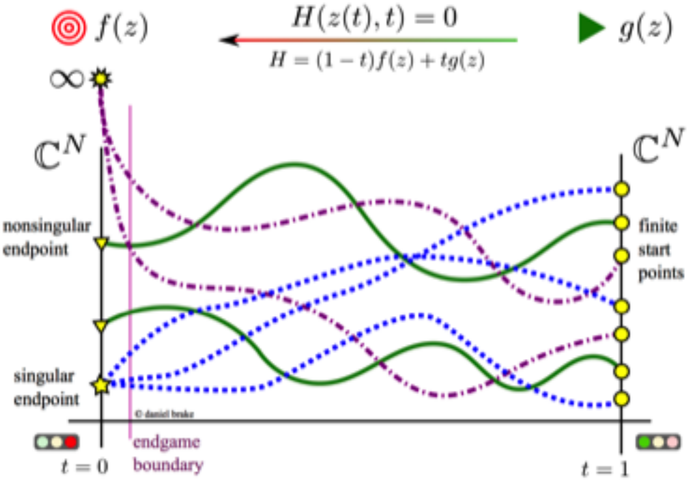


Figure: Homotopy Continuation Illustration (Dani Brake)

Homotopy Tracking

Theorem (Likelihood Geometry Group)

Fix a generic data vector u with positive entries. Let c_{win} and c_{stat} be scalings with positive entries.

Homotopy Tracking

Theorem (Likelihood Geometry Group)

Fix a generic data vector u with positive entries. Let c_{win} *and* c_{stat} *be scalings with positive entries.Consider the homotopy with target system the critical likelihood equations for the model* $V_A^{c_{stat}}$ *and start system the ones for* $V_A^{c_{win}}$*, with data vector u.*

Homotopy Tracking

Theorem (Likelihood Geometry Group)

Fix a generic data vector u with positive entries. Let c_{win} and c_{stat} be scalings with positive entries.Consider the homotopy with target system the critical likelihood equations for the model $V_A^{c_{stat}}$ and start system the ones for $V_A^{c_{win}}$, with data vector u.
Let $\hat{\theta}_{win}$ and $\hat{\theta}_{stat}$ be the respective MLEs and let γ denote the path of the homotopy whose start point (at $t = 1$) corresponds to $\hat{\theta}_{win}$.

Homotopy Tracking

Theorem (Likelihood Geometry Group)

Fix a generic data vector u with positive entries. Let c_{win} and c_{stat} be scalings with positive entries.Consider the homotopy with target system the critical likelihood equations for the model $V_A^{c_{stat}}$ and start system the ones for $V_A^{c_{win}}$, with data vector u.
Let $\hat{\theta}_{win}$ and $\hat{\theta}_{stat}$ be the respective MLEs and let γ denote the path of the homotopy whose start point (at $t = 1$) corresponds to $\hat{\theta}_{win}$. Then, the endpoint of γ (at $t = 0$) is $\hat{\theta}_{stat}$.

Proof Sketch

- By *Birch's Theorem*, a homotopy between the two systems is given by

$$H(\theta, t) := t\left(A\hat{p}_{stat} - \frac{1}{N}Au\right) + (1-t)\left(A\hat{p}_{win} - \frac{1}{N}Au\right)$$

Proof Sketch

- By *Birch's Theorem*, a homotopy between the two systems is given by

$$H(\theta, t) := t\left(A\hat{p}_{stat} - \frac{1}{N}Au\right) + (1-t)\left(A\hat{p}_{win} - \frac{1}{N}Au\right)$$

- This simplifies to $A \cdot (\hat{p}_{c(t)} - \frac{1}{N}u)$ where $c(t) = tc_{stat} + (1-t)c_{win}$

Proof Sketch

- By *Birch's Theorem*, a homotopy between the two systems is given by

$$H(\theta, t) := t\left(A\hat{p}_{stat} - \frac{1}{N}Au\right) + (1-t)\left(A\hat{p}_{win} - \frac{1}{N}Au\right)$$

- This simplifies to $A \cdot (\hat{p}_{c(t)} - \frac{1}{N}u)$ where $c(t) = tc_{stat} + (1-t)c_{win}$
- For positive real c_{win}, c_{stat}, we have $c(t) > 0$ for any $t \in [0,1]$. Thus by *Birch's Theorem* there is exactly one positive real solution to the system at every point along the homotopy path.

Proof Sketch

- By *Birch's Theorem*, a homotopy between the two systems is given by

$$H(\theta, t) := t\left(A\hat{p}_{stat} - \frac{1}{N}Au\right) + (1-t)\left(A\hat{p}_{win} - \frac{1}{N}Au\right)$$

- This simplifies to $A \cdot (\hat{p}_{c(t)} - \frac{1}{N}u)$ where $c(t) = tc_{stat} + (1-t)c_{win}$
- For positive real c_{win}, c_{stat}, we have $c(t) > 0$ for any $t \in [0,1]$. Thus by *Birch's Theorem* there is exactly one positive real solution to the system at every point along the homotopy path.
- Left to show tracking paths do not intersect (we show the Jacobian matrix of the system has always full rank)

□

Possible Applications

- In practice, a statistical toric model will come with a specified scaling c_{stat}.

Possible Applications

- In practice, a statistical toric model will come with a specified scaling c_{stat}.
- Knowing how scaling vectors c affect the ML degree of a particular toric model V_A allows us to find a convenient c_{win} (e.g. such that the model has *low* ML degree).

Possible Applications

- In practice, a statistical toric model will come with a specified scaling c_{stat}.
- Knowing how scaling vectors c affect the ML degree of a particular toric model V_A allows us to find a convenient c_{win} (e.g. such that the model has *low* ML degree).
- By the Theorem, we can now find the MLE $\hat{\theta}_{win}$ and track its unique homotopy path to find the original MLE of interest $\hat{\theta}_{stat}$.

Application Example

Example (Veronese revisited)

Recall
$$A = \begin{bmatrix} 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix},$$
with $u = (1, 3, 5, 7, 9, 2)$. Here $c_{stat} = (1, 1, 1, 1, 1, 1)$.

Application Example

Example (Veronese revisited)

Recall

$$A = \begin{bmatrix} 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix},$$

with $u = (1, 3, 5, 7, 9, 2)$. Here $c_{stat} = (1, 1, 1, 1, 1, 1)$. By choosing $c_{win} = (1, 2, 1, 2, 2, 1)$, the ML degree drops to $\mathbf{1}$.

Application Example

Example (Veronese revisited)

Recall

$$A = \begin{bmatrix} 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix},$$

with $u = (1, 3, 5, 7, 9, 2)$. Here $c_{stat} = (1, 1, 1, 1, 1, 1)$. By choosing $c_{win} = (1, 2, 1, 2, 2, 1)$, the ML degree drops to **1**. Computing the unique critical point we obtain the MLE $\hat{\theta}_{win} = (.0493827, 1.83333, 1.66667)$. Tracking this point in the homotopy we arrive at the point $\hat{\theta}_{track} = (.0863377, 1.63267, 1.51507)$.

Application Example

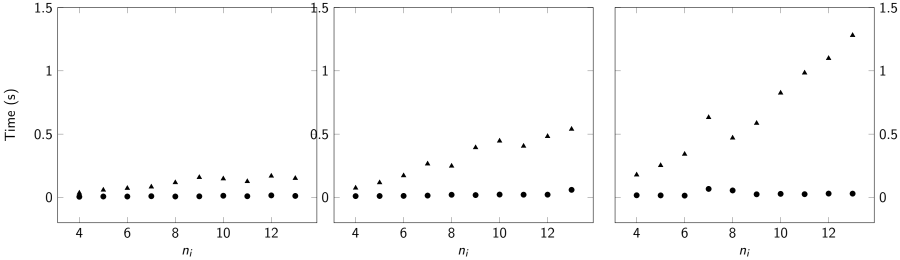
Example (Veronese revisited)

Recall

$$A = \begin{bmatrix} 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix},$$

with $u = (1, 3, 5, 7, 9, 2)$. Here $c_{stat} = (1, 1, 1, 1, 1, 1)$. By choosing $c_{win} = (1, 2, 1, 2, 2, 1)$, the ML degree drops to $\mathbf{1}$. Computing the unique critical point we obtain the MLE $\hat{\theta}_{win} = (.0493827, 1.83333, 1.66667)$. Tracking this point in the homotopy we arrive at the point $\hat{\theta}_{track} = (.0863377, 1.63267, 1.51507)$. This coincides with the MLE $\hat{\theta}_{stat}$ computed before.

Success story



$d - 1 = 5$ $\quad$ $d - 1 = 10$ $\quad$ $d - 1 = 15$

Figure: Running times of iterative proportional scaling (triangles) versus path tracking (circles) on rational normal scrolls. Average of 7 trials.